Faring with Facets:
Building and Using Databases of Student Misconceptions

Tara Madhyastha* and Steven Tanimoto**

*Facet Innovations, LLC.	**University of Washington
1314 N.E. 43rd St., Suite 207	Dept. Computer Sci. & Engineering.
Seattle WA 98105	Box 352350, Seattle, WA 98195
USA	USA
www.facetinnovations.com	www.cs.washington.edu

Abstract: A number of educational researchers have developed pedagogical approaches that involve the teacher in discovering and helping to correct misconceptions that students bring to their study of their subject matter. During the last decade, several computer systems have been developed to support teaching and learning using this kind of approach. A central conceptual construct used by these systems is the "facet" of understanding: an atomic diagnosable unit of belief. A formidable challenge to applying such pedagogical approaches to new topic areas is the task of discovering and organizing the facets for the new subject area. This paper presents a taxonomy of misconceptions and a methodology for going about the task of preparing a database of facets. Important issues include the generality and diagnosability of facets, granularity of facets, and their placement on a scale of problematicity. Examples are drawn from the subjects of physics and computer science and in the context of two computer systems: the Diagnoser and INFACT.

Keywords: educational assessment, misconception, teaching, facet, ontology, concept, pedagogical knowledge, facetbase, diagnosis, preconception, learning environment, computer-assisted instruction.

1 Introduction

1.1 Motivation

Students do not come to the classroom as blank slates; they have prior knowledge. Because learning involves transferring this knowledge to new situations, prior knowledge can actually make it difficult to learn new things. Bransford, Brown et al. (1999) give a variety of examples of this phenomenon in multiple domains.

There has been much research, especially in the area of physics, on specific student misconceptions that interfere with learning. An example is McCloskey (1983). These are often referred to as "preconceptions" or "alternative conceptions," but they have in common that they are conceptions that are different from the optimal understanding and may interfere with optimal understanding of the target concept, as suggested by Hammer (1996). For example, students who believe that a moving object has an "impetus" that propels it have a difficult time accepting a Newtonian theory of forces as interactions.

The analogy to medicine – that you can identify and "treat" student conceptions that interfere with learning with efficient, targeted interventions, is alluring. In reality, we do not know the best way to do this, and the literature is filled with conflicting results. However, even if identification of student prior knowledge is not actually the most efficient way to help a student to learn from a cognitive perspective, it may have a huge impact on student engagement. A student who is excited by discussing his or her ideas may become more motivated to learn.

This creates a need to identify student preconceptions and use them to guide classroom activities. Such a strategy is one way to perform and use formative assessment, which has been shown by many, such as Black and Wiliam (1998), to raise standards. There are many ways in which understanding of student preconceptions may be used to improve student learning. For example, highly motivated students may be able to receive individualized prescriptive lessons that they could do as homework on their own. A teacher might choose to address the two most common problems in the classroom with guided activities.

1.2 Facets

When students learn, their knowledge is fragmented, often inconsistent, and sometimes incorrect. A "facet" is an attempt to categorize these partial understandings in a way that can be communicated to a teacher. From the point of view of a teacher, facets are context-sensitive fragments of understanding that students can demonstrate through their answers to diagnostic multiple-choice questions, hand-coding or automated analysis of text, or through a Socratic dialogue with a

student in a classroom or online. According to Minstrell, facets are "slight generalizations from what students actually say or do in the classroom" (Minstrell 2001). This understanding may be correct (a goal facet), incorrect and self-consistent (a misconception), or reflect a level of mastery of a subject.

1.3   Examples

Before discussing the details of facets and their construction, we present two examples to clarify the notions and context for that discussion. In order to explain our notion of "facet" we re-introduce an example used by McCloskey, Minstrell and others, from the subject of physics. Then, in order to show the use of facets in the context of a computer-based learning environment, we present a facet from the subject of image processing as taught at the college-freshman level.

1.3.1. Example from Physics

In an introductory physics class, students typically come into the classroom with definite preconceptions about physical phenomena, as argued by McCloskey (1983). An example from kinematics is the notion that moving objects, unless continuously powered, will eventually come to rest. This notion is at odds with Newton's first law of motion, which says

> *Every object in a state of uniform motion tends to remain in that state of motion*
> *unless an external force is applied to it.*

The misconception can be blamed on the fact that on earth (and in space, but less so), essentially all moving bodies encounter a slowing force known as friction. Whereas the motive force propelling an object is typically visible (e.g., the baseball pitcher's arm) or the invisible but well-known force of gravity, frictional forces do not show the same kind of evidence of activity. Students therefore sometimes don't realize that friction is an "external force" on the moving object.

This misconception is a facet of understanding motion. The term "facet" makes an analogy to the shape of gemstones. By seeing one facet of a concept, a student has a partial conceptualization. In the physics example, the student indeed has intuition about the motion of objects, but a part of the concept is missing: the notion of friction as an external slowing force.

The value of diagnosing a student's misconception lies in the possibility to offer particular, efficient instruction that takes the student from their current cognitive state to one embodying a full and correct understanding of the target concept. For the motion facet above, such instruction would typically involve introducing the related concept of friction and challenging the student to construct the relationship between friction and moving objects that slow down.

1.3.2 Example from Image Processing

Let's now consider a different kind of example. In this case the domain is nominally image processing. However, the facet we describe really relates to understanding the mathematics of functions, something that is officially covered in the high-school mathematics curriculum. A fundamental difference between the physics example and the one we are about to present is that students experience motion as a phenomenon in everyday life, and they therefore have definite preconceptions about it, but in the mathematical domain, any preconceptions they have tend to be the result of earlier study or they tend to derive from analogies.

A fundamental kind of activity in image processing is to transform images. For example, starting with a digital image from a modern digital camera, the image can be brightened by increasing the brightness of each pixel. The result is a new image. Another common transformation is to rotate an image 90 degrees clockwise (or counter-clockwise); this is a typical chore for amateur photographers nowadays, as they upload their pictures to a web site or PC photo album. Rotation (by any angle) is an example of a geometric transformation.

An image processing system applies a transformation to an image by computing, for each pixel of the output image, an appropriate colour value, based on the values of pixels in the input image. There are two ways one can imagine this being performed. One way starts with a pixel in the input image and figures out where to put it in the output image. This is called the "push" method, because the pixel values of the original image are sent or "pushed" to the other (the "range") image. However, image processing systems, with few exceptions, do not use the push method. Doing so would typically result in unnecessary "holes" in the output image – pixels where no data from the input image was sent. The alternative method, called the "pull" method, figures out, for each pixel of the output image, what value (or combination of values) from the input image should be taken, and these values are "pulled" from the input image into their places. Although with the pull method there could still be pixels with unspecified values, they tend to occur at the borders of the image rather than as the numerous, small holes that often occur with the push method.

Even after an explanation of how an image processing system works, it is common for students to exhibit a "push-method" facet, when they should exhibit the (correct) "pull-method" facet. The explanation for this comes directly from

the mathematics involved. Let's consider the problem of devising a formula that will have the effect of shifting the image 5 pixels to the right.   Assume that we are working with a monochrome (grey-scale) image; that is simpler than colour, because it means that we only have to compute one value per pixel. Assume also that our formula will be in terms of two variables: x and y, and that they refer to the horizontal and vertical coordinates of the pixel we wish to compute. The original image is represented by a function S of two coordinate variables – we could call them u and v to distinguish them from the previously mentioned x and y, but rather than using u and v, we'll use expressions to indicate the coordinates of the desired image pixel. The correct formula for shifting the image 5 pixels to the right is this:

$$S(x - 5, y)$$

This means that the output pixel value for (x, y) should be taken from the source image at a position 5 pixels to the left of the point (x, y).

If a student is asked to give a formula that shifts the image 5 pixels to the right, instead of the above formula, the student might give the formula

$$S(x + 5, y)$$

Then this is evidence that the student holds the "push-method" facet. The student's logic typically is "we want the pixels to move to the right, and so we have to increase the x value."

The same general concept and misconception arise in high-school mathematics, typically in the coverage of parabolas in analytic geometry. The general formula for a quadratic equation in "analyzed" form is $(x - h)^2 - 4p(y - k) = 0$.

The parameter h is the horizontal position of the vertex of the parabola. (Also, the vertical position is given by k and the distance of the vertex from the focus is given by p.) If we wish to alter the formula so that it represents a parabola shifted 5 units to the right, we must increase h by 5 (which means subtract 5 more units from x) within the parentheses. This subtraction seems counterintuitive to students who have the push-method facet.

The facets for a concept form a group that we call a "cluster," borrowing from the terminology of Minstrell. The number of facets in a cluster typically varies between two and six. The cluster in the image processing facet base containing the push and pull facets is illustrated in Figure 1. This figure is a screen shot from the facetbase display tool within the INFACT system, an online learning environment supporting facet-based diagnosis. The INFACT system is described by Tanimoto, Carlson, Husted, Hunt, Larsson, Madigan, and Minstrell (2002).

Cluster "Basic (Push-Pull)"

A Cluster in Subject "PixelMath Operation" in the Infact Facetbase pixelmath_2003

This facetbase is accessible to the forum "Pixels Class (Winter 2004)" and possibly others.

Cluster Explanation

How PixelMath's image processing engine computes values using formula symbols x and y that refer to the coordinates of the CURRENT DESTINATION image pixel.

Level	Facet Name	Explanation
0	Understands the "pull" method	Realizes how x and y in a formula are bound to the current destination pixel coordinate values during the computation of a destination pixel's R, G, and/or B components.
4	Thinks "push" rather than "pull"	Operates under the assumption that x and y refer to source image locations, even for transformations involving arithmetic on the coordinates, such as Source1(2x, 2y), which actually shrinks the image rather than enlarge it. If a student reports that this formula should enlarge the image, then this is clear evidence for this facet.
6	No clear concept of either push or pull	Exhibits enough inconsistency or uncertainty in applying coordinates in transformations to suggest the student has no strong conception of either pull or push methods.

Figure 1. A portion of the facet base for introductory image processing. The cluster for the push-pull distinction is shown. There are three facets in the cluster. They have been assigned levels 0 (expert), 4 (somewhat problematical), and 6 (weak).

Diagnosing a facet such as the pull-method facet can take several approaches. A direct approach is to ask the student to write down a formula that is supposed to have a given effect, such as to shift an image 5 pixels to the right or to magnify the lower-left quarter of an image by a factor of 2 in each direction. If the student gives a formula with the wrong operation (i.e., minus for plus or multiplication for division), that is taken as evidence that the student holds the push facet rather than the pull facet. An alternative approach is to present the student with a formula and ask for a prediction of what it will do to the image. Even less direct is to engage students in group discussions about transforming images and find within their conversations, evidence of facets similar to the evidence in the first two approaches.

Once a student has been diagnosed with the push facet (the incorrect one), for example, a possible intervention is to ask the student to work through an exercise sheet that contrasts the push and pull methods. On the other hand, a student who demonstrates holding the pull facet might be asked to explain the formula to a student having the push facet.

The example from physics and the one we've just given from image processing illustrate an important aspect of the facet of understanding. A misconception usually contains within it some part or parts of the correct conception. If the instructor can correctly diagnose the facet, then it's only necessary to teach the missing components – to fix the part that's broken or incomplete. How broadly can this methodology be applied? Are there suitable facets in other subject areas that will permit instructors to gain such efficiency in teaching?

There are many ways of cataloguing student preconceptions. In the next section we consider student conceptions broadly and relate them to the facet notion.

2 Taxonomy of Student Conceptions

2.1 Background

Broadly speaking, there are two theoretical perspectives describing how students' knowledge is organized. The first camp is the "knowledge-as-theory" camp, which holds that students form naïve but unified and coherent frameworks of knowledge. The second camp holds that students have only loose ecologies of ideas, with little consistency. These different theoretical perspectives suggest different levels of diagnosis and types of intervention. For example, a strong theoretical misconception might be challenged by hypothesis testing, but this approach might not be as successful for an inconsistent collection of ideas.

Misconceptions may be seen as a way to categorize some ideas from the "knowledge-as-theory" camp. A teacher cannot help a student past their misconceptions without first challenging them. Thus, a misconception might be thought of as a diagnosis for an erroneous student belief that can be "treated" with some kind of appropriate instructional intervention.

One difficulty with misconceptions is that often they are context-sensitive, and therefore unstable; certain contexts might trigger certain misconceptions and others do not. Students might not have consistent underlying models. Moreover, a misconception in one context might not be problematic in another context. To address this, diSessa (1985) proposed an alternative way of viewing student knowledge called phenomenological primitives (p-prims), which are very general reasoning strategies that can be activated or not depending on context. P-prims are an underlying construct that describe loose ecologies of ideas. An often-cited example is Ohm's p-prim, which encapsulates the idea that more effort implies more result. This is a fundamental idea that is based on so much experience that it is difficult to identify why one believes it. For example, a student might use this p-prim to state, on a formative assessment, that when a large truck hits a small car, the larger truck must have exerted a larger force because it is less damaged. Recognizing that this p-prim is at work, a teacher may choose to probe the difference between the equal forces and the actual reasons for the resulting interaction.

Minstrell also departed from the misconceptions research to describe "facets," which are "slight generalizations from what students actually say or do in the classroom" (Minstrell 2001). These generalizations provide a common language to describe student ideas that helps students, researchers, and teachers to communicate. Facets are context-sensitive fragments of understanding that students can demonstrate through their answers to diagnostic multiple-choice questions, hand-coding or automated analysis of text, or through a Socratic dialogue with a student in a classroom or online (Hunt and Pellegrino 2002).

One might relate facets to diSessa's research by noting that a facet is the result of an application of a p-prim to a particular problem context. Minstrell *et al*. have catalogued a large number of facets exhibited by students in middle school physics in a variety of topics. Each topic, or "facet cluster" has approximately 10-20 facets, coded loosely numerically according to how problematic they are (more problematic facets generally require more instructional effort to bring the student to optimal understanding). All facets, including "goal facets," are organized within this structure. Using an online tool called Diagnoser, a teacher would identify the facets most prevalent in the classroom and, based on the facet diagnosis, conduct "prescriptive activities" with the students. This method was described by Hunt and Minstrell (1994) and Minstrell (2000).

From the classroom perspective, a teacher who can identify student misconceptions can address them. A teacher who sees a palette of facets can treat the problematic ones as specific pedagogical approach that is most effective is a matter of research; however, diagnostic information is required to make the pedagogical decisions.

Although diagnostic information of the sort described above seems at odds with the kind of dimensional ability information available from traditional tests, such as standardized exams, many facets might be loosely aligned to ability as measured by an achievement test in the topic area. This has been found in the areas of chemistry by Scalise and Wilson (2005) and for many facets describing student understanding of motion. In an area without a large body of research on the kinds of mistakes students make and how they acquire knowledge, a facet might be as simple as whether a student knows the answer or not. In these domains where student responses are not governed by a rich set of p-prims resulting in a complex facet-base, it makes more sense to talk about mastery levels. This assumes that the resulting intervention looks less like an experiment designed to target some persistent misconception and more like a dialog with the student suggesting areas for increased study.

2.2 Taxonomy

One might attempt to diagnose conceptions through a variety of strategies at these levels (which are loosely ordered from first to last according to the sophistication and consistency of reasoning applied). Note that the purpose of making a diagnosis is to allow a match between a diagnosis and an appropriate instructional intervention. For each category, we

describe the kinds of interventions that may be applied. We also give an example of a diagnosis using the problem of determining the speed of an object from a graph.

Levels of mastery: At the base level, a diagnosis may simply record the student's level of mastery of a subject, according to a pre-determined scale. This kind of information tells an instructor what a student knows, and how well, but not why they might be having difficulties. As such, it best applied to a topic area where there are not many preconceptions that might interfere with acquiring knowledge. Suitable interventions may involve doing additional work to help the student learn the material.

To obtain level of mastery level about a student's ability to determine the speed of an object, one might pose several questions of increasing difficulty. For example, it is easier to determine the speed of an object at a particular time from a speed vs. time graph than from a position vs. time graph. The latter involves making a calculation of slope.

Partial Conceptions: Levels of mastery are concerned primarily with measurement of the correct ideas held by students. To learn slightly more about the incorrect ideas they might hold, one can diagnose partial conceptions. When a student does not completely understand something, partial conceptions should lend insight into what beliefs they have. This category is particularly useful in topic areas where students have a variety of preconceptions that may interfere with learning.

For example, when first learning about graphs, students often use other primitive strategies, such as treating a graph as a map of motion. This might work when a student examines a graph of displacement over time, but not when viewing a graph of speed over time. To determine if a student has this incorrect idea, one might ask the student to view a graph showing a reduction of speed over time and describe the motion of the object. A response of the form "it is rolling downhill" might indicate this facet.

Misconceptions: Although facets give more diagnostic information than levels of mastery, they may not be indicative of a strongly-held reasoning strategy. As such, it is overkill to address a facet if it cannot be diagnosed reliably. In some cases, students have very strongly-held misconceptions that must be addressed with very specific kinds of interventions. In some situations, it may be easier for a teacher to focus on recognizing and addressing these misconceptions. Misconceptions are therefore a subset of facets, with less breadth and applicability. For example, a classic misconception held by even college students is that the seasons are caused because of the changing distance from the sun to the earth. Despite the fact that this explanation is easily challenged, it is simple, and employs the idea that "closer to a heat source is warmer". A teacher might lead a student through structured activities to break down this misconception while engaging their beliefs.

An example of a misconception regarding speed of objects is the idea that an object cannot be moving at a single instant in time. Many students who do not have the concept of a derivative reason that there is no such thing as instantaneous motion. A question that asks students the speed of an object at a particular instant can elicit this misconception.

p-prims: diSessa's p-prims move beyond observable knowledge states to the ways in which students apply their experiences to govern acquisition of new knowledge. Diagnosis of a p-prim is context-sensitive, since students might apply different reasoning in different circumstances. It may not map directly to a specific educational intervention, rather, it might help to understand what "tools" a student is applying to the learning process. However, p-prims may be viewed as the underlying agents resulting in misconceptions and facets.

For example, a student may describe the speed of an object on a position-time graph as slowing down if the line slopes upward, because it is going uphill, invoking a p-prim that the graph is a map of motion. The same student might respond that the object is slowing down if the line slopes downward, because higher (on the graph) means more speed. This is an example of using different p-prims to solve this problem. Regardless of the specific reasoning used, the student has difficulty understanding graphical representations, and needs practice drawing graphs and connecting those graphs to actual motion.

Patterns of thought: At the highest level, we might attempt to diagnose strategies of reasoning applied in different contexts. In the language of p-prims, why does someone apply one or the other in particular contexts? When asked to solve a problem using a simulation, does a student begin with prior knowledge to zero in on the correct solution, or do they attempt to solve it systematically (or unsystematically) This kind of information might be used to help a student move to more efficient problem-solving strategies.

A pattern of thought might be diagnosed by identifying unique patterns of response to the kinds of questions described above.

3   Creating a Facet Catalogue

Once you have identified the level, or levels, of the taxonomy which you would like to diagnose, you begin to design and build a facetbase. This section discusses the process of designing and building a facetbase for use in teaching or educational research. We use the term "designer" to refer to the author of the facetbase. This person is presumably a teacher, an educational materials developer, or an educational researcher. The section begins with six questions the designer should answer before taking any additional steps. Then, it presents several methodologies for building facetbases, and discusses how to start using draft facets that we call "proto-facets." Next is a discussion of how to cover a concept with one or more clusters of facets. Finally, we describe how to assign a numeric "problematicity" value to each facet that helps to position the facet within its cluster.

3.1 General Questions to be Answered

We begin with six general questions, plus additional questions whose answers may help to answer the general questions.

Purpose: What is the purpose of the facetbase and level of accuracy intended in its coverage? Will the facetbase be used simply as an organizational aid to the teacher? Will it serve as a framework for automated assessment through online testing, rule-based feedback, etc? Will it be used to help report progress to students, parents, or others? Will it be the basis for personalized instruction or suggestions to students?

Process: What process is intended for drafting, refining, validating, and maintaining the facetbase? How much time is available before the facetbase must be ready for its intended uses? Will it be the work of one person or of a team? What might be a realistic timeline for the various development and testing stages?

Scope: What is the intended scope of the facetbase in terms of subject content? What list of concepts is to be covered? Is there a specific context for these concepts? Surface features of questions or problems can influence student responses. What contexts are most important?

Granularity: To what degree of granularity will concepts and misconceptions be represented? Roughly how many misconceptions are expected per concept?

Sources: What materials and other sources of information are available to the designer? More particularly, which of the following are available?

- textbooks (printed or electronic),

- student expression (text bases or recordings),

- curriculum specialists,

- master teachers,

- pedagogical guides and teaching literature.

Methodology: What methodology will be used to develop the facetbase? Several methodologies are described in the next section.

3.2. Development Methodologies

There are many approaches to developing a catalogue of conceptions and misconceptions about a subject area. Here we describe three. These are not mutually exclusive, and they can therefore be combined in whatever ways most suit the designer.

A *top-down* approach begins with the subject area and the learning objectives of the curriculum, and breaks that down using the pedagogical content knowledge of the field into clusters (sub-topics) and within those, specific facets. These facets may be at any level of our proposed taxonomy, ranging from knowledge states to patterns of thought.

Conversely, a *bottom-up* approach begins with student conceptions (for example, as noted in a classroom or through classification of responses to open-ended assessments) and uses the student-mined facets to form the facetbase, organizing them as appropriate.

Ultimately, a top-down approach will require validation that the students, in fact, have facets corresponding to those predicted. Likewise, a bottom-up approach ideally incorporates some prior pedagogical content knowledge to determine

what kinds of assessments might elicit problematic ideas. Thus, a compromise approach is *middle-out*: translation of pedagogical expertise into facetbase form. Here, one uses prior research into student misconceptions in a particular domain, together with pedagogical expertise, to identify clusters. Then the designer creates open-ended questions that are designed to elicit problematic student ideas and administers them to a representative group of students. The responses are then mined to determine specific facets that are then organized into clusters. This process is iterative, until the designer has settled upon a set of facets that serve the desired instructional use.

3.3. Proto-facets

A "proto-facet" is an easily created, facet-like category that can serve as a draft facet, later to be either retained as-is, refined, or eliminated. Because they are easy to create, they are ideal starting representations in the development of a facetbase. Their descriptions can typically be generated by cutting and pasting from other materials or by simple paraphrasing or summarizing.

Here are five kinds of proto-facets. The first three support the top-down methodology of development, while the fourth supports the bottom-up methodology. The fifth kind is bottom-up in a different way, closely mirroring a set of possible instructional interventions.

1. textbook definitions and other statements of "truth";

2. negations of textbook statements;

3. the null facet: a student has "no clue" --- the student has no relevant idea;

4. raw student expressions, statements taken directly from student responses;

5. descriptions of particular instructional interventions or learning materials.

The first type of proto-facet will normally serve as the expert-like facet within a single cluster. The second type serves as a catch-all for misconceptions, while the third type represents straight ignorance.

The fourth type of proto-facet is used to capture both a conception or misconception and actual evidence for it. ("I think heavier objects always fall faster than lighter objects.") A small amount of generalization can be performed by the designer so that the proto-facet handles some alternative expressions of the same idea. (e.g., "[heavier | bigger] [objects | things] fall [faster | quicker | more quickly] than [lighter | smaller] [objects | things | ones]"). However, generalization beyond superficial language variations would take us away from proto-facets into more refined candidate facets.

3.4. Structure of a Facet Catalogue

In addition to the facets themselves, a facet catalogue requires an organization. Let's consider one possible organization -- that used in the INFACT system. Within INFACT, a facet catalogue is known as a "facetbase." The hierarchy within an INFACT facetbase has four identifiable levels.

Top-level object: The facetbase is a named collection of subjects, clusters, and facets. It is named with an identifier legal within the Unix operating system as a file name. It generally resides on a particular server computer. There may be several facetbases on the same computer.

Subject: There may be any number of subjects within a facetbase. A subject has a name and a description as well as an author id (the user number for the person who created the subject).

Cluster: There may be any number of clusters within a subject. Like a subject, a cluster has a number, description, and author ID number.

Facet: There may be any number of facets in a cluster, but normally between 2 and 10. Like a cluster, a facet has a name, a description and an author id. It also has a problematicity value in the range 0 to 9. These values and how to assign them are described later in this section.

3.5. Covering a Concept

Given a concept that is to be represented in the facetbase, there are several steps to take to design and complete its

representation. Both the correct conception and the associated misconceptions need to be taken into account.

3.5.1. Analyzing the Concept and Its Misconceptions

The first few steps involve analyzing the concept:

A. Enumerating important aspects of the concept. If there are component subconcepts, these should be identified and their relationships to the main concept written down. If there are alternative manifestations of the concept that students are likely to be acquainted with, these should also be written down.

B. Deciding whether to set up one cluster for the concept or a separate cluster for each aspect of the concept. This decision may be influenced by any of the following: (1) extent to which a clear subdivision of the concept into subconcepts is available, (2) the expected level of effort in diagnosing all the subconcepts rather than simply the overall concept, and (3) likely availability of different interventions that are appropriately adapted to the various student facet profiles that could result. If only a few alternative interventions are expected to be available, then there would seem to be relatively less value in modelling and diagnosing students' understanding at the more detailed level.

C. Choosing a schema for the (each) cluster. Here are five prototypical schemata for organizing the set of facets within a cluster.

(1) the binary cluster schema. This schema provides for only two facets in the cluster: an expert-level facet, and another, catch-all facet for all other cases.
(2) the ternary cluster schema. Slightly more refined than the binary cluster, this cluster includes the same expert-level facet but splits the remaining cases into those involving some kind of misconception and those that correspond essentially to a state of ignorance.
(3) the quasi-continuous cluster schema. In this schema, a clear scale of levels of expertness is recognized, and each facet is closely associated, if not defined, with a position on this scale. Most facets in the DIAGNOSER system fall into this schema.
(4) the power-set cluster schema. This schema acknowledges the existence and importance of subconcepts or essential aspects of the main concept, and it establishes facets in correspondence with some or all subsets of these aspects. A full power-set schema provides a facet for every subset. However, if there are more than 3 or 4 such essential aspects, then the power set becomes awkward to manage because of it size, and it becomes less and less likely that each element of the power set will find a student with the corresponding subset of aspects. However, one variation of the power-set cluster schema creates a few facets for those subsets likely to correspond to student cognitive states, and it may lump multiple subsets into a single facet to achieve economy in the facetbase.
(5) the mutually-exclusive, qualitatively distinct schema or "bunch" cluster schema. This schema recognizes the existence of distinct beliefs, where any given student can only hold one of the beliefs at a time.
D. Describing the schema, especially if it is not one of the "standard" ones discussed previously.

E. Enumerating the facets within the (each) cluster. This means categorizing each facet according to the cluster to which it belongs.

3.5.2. Creating a Facetbase Entry

After these higher-level decisions have been made, there are detailed steps to be taken in order to finish the facetbase entry for a concept:

1. Naming the facet. The name should naturally be unique within the cluster, and for simplicity, perhaps should be unique within the subject or within the facetbase, as well.

2. Giving an explanation of the facet. The explanation is intended for teachers who use the system in the course of teaching.

3. Writing a hint for diagnosing the facet. Such a hint may identify typical suggestive evidence for the facet.

4. Explaining any important relationships between this facet and others, except those implied by the schema. For example, if there is another facet that could easily be confused with the current one, this should be mentioned and the distinction explained.

These various items can be recorded using computational facetbase editing tools. One such tool is described later.

3.6. Problematicity of Facets

While facets could be used in a purely qualitative fashion, it is usually possible to draw distinctions among them in terms of how expert-like or how problematical they are. There are several reasons for assigning numeric values to the facets within each cluster. The most important is so that a teacher can readily judge how far a student needs to go in order to reach the expert level. Another is so that the facets can be listed in a systematic order. A third is so that there is a ready basis for showing a student's conceptual progress over time. Numeric levels for facets are used by the INFACT software in plotting trends and making aggregate reports.

The aspect of facets used for ranking is one we call "problematicity." Here are some operational definitions of problematicity: (1) A facet is "problematic" if the student just cannot be permitted to complete the course holding that facet. (2) A facet is "problematic" if holding this misconception blocks acquisition of some other essential concept. The idea of "problematicity" takes on different meanings based on the kind of diagnosis we make. In essence, we are trying to measure something that falls into the taxonomy of facets described in Section 2.2. Many tests are developed using Item Response Theory, which models a particular student's performance along one or more axes of achievement according to their responses to a set of questions of varying difficulty. IRT models are limited by the underlying assumption that students are generally similar, meaning that the relative order of difficulty of items is the same for any two students. Nevertheless, it is a good fit for many circumstances. A level of mastery is not very different from an ability level as determined by an IRT model. In this sense, "problematicity" is useful in terms of measuring the "distance" of the student from expert understanding.

Facets may also contain information about why a student does not understand something. In many cases, the problems students have are similar, and may be ordered according to a progression of expertise, or ability level. Some facets, however, are extremely persistent (misconceptions). In this case, "problematicity" may take on a new meaning: how hard is it to move a student from their deeply-held misconception to a more correct understanding? Regardless of the expertise level associated with such facets, the learning and instruction required to overcome them is disproportionate to their proximity to expert understanding. In this sense, misconceptions are "the most problematic" facets, in terms of the difficulty to change students' minds.

Higher-order kinds of facets, such as p-prims and patterns of thought, can be rated by problematicity only when viewed in context. Thus, although it may be useful to make a diagnosis at that level for pedagogic reasons, problematicity can be judged only in terms of the facets or misconceptions that a student manifests.

3.7. Assigning Problematicity Values to Facets

After identifying the facets in a cluster and ranking them in terms of problematicity, the next step is to place them on a numerical scale, such as one that goes from 0, representing the most expert-like end of the spectrum, to 9, representing the most problematical end of the spectrum.   This problematicity scale then allows rough but useful distinctions to be made. A value of 0 means that the facet represents a full and accurate understanding of the concept (expert-level). A value of 9 means that the facet represents a seriously flawed understanding: one that may be worse than knowing nothing at all about the concept.

The INFACT software requires that integer values in the range 0 to 9 be used for expertise levels. This limited range and precision of values is in part an acknowledgement of the inherently inaccurate and subjective nature of assigning numerical values to misconceptions at all. It should also remind teachers that it is important to provide qualitative comments to students, ideally facet-oriented feedback, whenever practical to do so.

One approach for assigning problematicity values to facets is to base the values on the schema used to organize the facets within the cluster. For example, if the cluster follows the binary schema, then use the values 0 and 5 (excellent vs. no clue)

Values for ternary clusters might be set at: 0, 5, and 8. In the case of a power-set schema, where the base set has either two or three elements, the assignment of values is likely to be based in part on the cardinality of the particular subset. However, the various elements of the base set may have different weights, according to their importance within the overall concept.

In the case of the schema of mutually exclusive, qualitatively distinct ideas, the problematicity values may have to be assigned on an ad hoc basis, respecting the relative degrees of problematicity for the ideas.

It is possible for two distinct facets within the same cluster to be given the same problematicity value. That would imply that any student who changes her or his conception from one of these facets to the other would neither be viewed as making progress nor regressing in understanding the concept.

3.8. Recording Meta-Data

The draft facetbase should not be considered complete without the inclusion of status information. Such meta-data should include the designer's assessment of the current state of development of the facetbase. What percentage of the concepts have been entered? How complete are the clusters and facets that correspond to them? What are the source materials used so far and where can copies of them be found? What steps still need to be taken in order to have a workable, testable version of the facetbase? What questionable decisions have been made and why? How certain is the designer about the correctness or accuracy of what has been done so far? What is the version number for this draft? When was it last modified? Who if anyone has proofread or reviewed what has been completed? Are any parts of this draft already in use, and if so, by whom and for what purpose?

3.9. Tools for Facetbase Constructions

To support the construction of facet bases, the INFACT project provides several tools within an overall system called INFACT. INFACT stands for Interactive Networked Facet-based Assessment Capture Tool. First of all INFACT-FORUM is a textual conferencing system (like a collection of newsgroups) in which students can do group work online and thereby create a corpus of writing that can be analyzed for assessment or facet-base construction purposes. INFACT-FACET-MANAGER is a facility that teachers and researchers can use to edit a facet base from any standard web browser.

4 Diagnosis

The true state of understanding held by a student is directly unknowable. However, we can diagnose understanding by posing questions and seeking evidence of particular ideas. The evidence may be the ideas themselves (e.g., an answer that reveals the student has a particular misconception) or a pattern of evidence that implies a higher-order reasoning strategy (that itself might be diagnosable). For example, Redish and Boa model the phenomenon that students are neither strictly Newtonian or Aristotelian in their reasoning by capturing their shifts in thought provoked by different contexts in multiple choice assessments; see Redish (2004). This pattern of shifting might itself be important in diagnosis and treatment.

4.1. Diagnosis in Medicine and in Education

The term "diagnosis" is based on analogy to a medical diagnosis. The implication is that it is important to learn something more about a student's thinking than whether they get an answer right or wrong to help them to a more expert understanding (i.e., treatment).

This analogy is particularly apt in the case where a student's preconceptions interfere with an understanding of what is being taught. A student will experience difficulty trying to reconcile the new information with what he/she already believes. A teacher must have more detailed information about the student's beliefs than a failed assessment to help (treat) this student.

However, medical diagnoses are often just statements describing the patient's condition with no information about why. Perhaps the reason is unknown, indeterminate, or simply irrelevant to the treatment. Similarly, facet diagnoses may simply be evidence of a student's knowledge state without information about why (e.g., the reasoning that led the student to this state). The most appropriate educational intervention might not require this detailed understanding.

4.2. How Diagnoses are Made in Diagnoser

Hunt and Minstrell (1996) developed the DIAGNOSER system based on the facet theory of Minstrell (1992, 2001) to be used by a teacher to diagnose student difficulties in science (www.diagnoser.com). The system consists of short sets of questions designed to elicit middle-school and high-school student thinking around specific concepts in physics. Each DIAGNOSER question is designed to elicit facets of student thinking that are then reflected in their multiple choice or numerical response. For example, Figure 2 shows the first question from a set on identifying forces. Each multiple choice distractor corresponds to a common facet that students exhibit in this context. Note that the context is designed to elicit a very specific diagnosis. As students take questions, they receive feedback about their thinking and reasoning.

Figure 2. Sample DIAGNOSER Question.

When students have completed sets, the teacher can view the diagnosed facets. For each individual student, the diagnosis consists of the list of facets corresponding to their multiple choice responses. If the student is asked to repeat a question, two facets are listed. For the class, the teacher can view each individual student's response pattern and a summary of the most frequently appearing facets. For each facet, teachers can view a description of what the student might be thinking, and recommended prescriptive activities.

4.3. Steps in Making a Facet Diagnosis in INFACT

The INFACT system supports diagnosis in essentially two different ways. One way is manual, and the other is automatic. Here we focus mainly on the manual procedure. (The automatic method requires writing and testing sets of rules, which is essentially a specialized form of computer programming.) The manual process contains the following steps.

- Deciding the purpose of diagnosis. Typical purposes of diagnosis include preparing to help a student overcome a conceptual problem; rating the progress of a student or a class; or evaluating the effectiveness of a particular educational activity or material.

- Choosing a level of effort. The practicalities of teaching involve careful management of time. Facet diagnosis tends to be time-consuming, and the accuracy of diagnoses can easily suffer if the teacher doing the diagnosis doesn't give enough time to reading what students write. If enough time is available, it might be possible to read through an archive of student writing in chronological order, and this might make it possible to follow threads of discussion among multiple students and correctly understand the intent of students in particular messages. On the other hand, if little time is available, it will be important to quickly get to the meat of discussions and base diagnoses on a small number of (hopefully) content-rich messages.

- Focusing on a particular cluster. The choice of cluster is most likely dictated by the importance of its concept in the curriculum. However, it may be based on the availability of student data for it: writing, event logs, etc.

- Identifying a "pregnant post." A message laden with facet-rich expression(s) is sometimes called a pregnant post. If the teacher has raised a conceptual question, then the direct answers to this question may be pregnant posts.

Also, certain keywords and phrases may be indicative of pregnant posts. "I believe that…" for example, suggests that a speculation or facet-rich expression is coming.

- Selecting the expression to serve as evidence. Within a message, there is typically one phrase or sentence that is suggestive of a particular facet. The rest of the message may consist of rhetorical material or problem-specific details. In this step, the teacher selects, typically by using the mouse, and highlights that portion of the message that directly supports the diagnosis.

- Selecting a facet. This diagnosis step consists of choosing, from a "facet browser," the facet most strongly indicated by the evidence.

- Creating a facet-assessment record and assigning a confidence value. The facet-assessment record is created simply by clicking on a button. This causes a database entry to be registered on the INFACT server. The confidence value is an integer in the range from 1 to 5 where 1 represents "very uncertain" and 5 represents "quite certain." This value can be useful when it comes time to act upon the diagnosis, discuss the student's progress, or evaluate the need for more diagnosis.

- Making a student-visible comment. In recognition of the reality of limited time on the part of teachers, INFACT provides a place in which a teacher can write a comment to the student related to the diagnosis. In general, the student does not see the diagnosis, but may see this comment. One version of INFACT allows sending this comment immediately in an email message.

Here is an example that comes from the use of INFACT in a course on image processing. One of the topics covered is the use of mathematical formulas to represent image transformations involving coordinate manipulations. The purpose of facet diagnosis in this case is to identify any particular difficulties a student may be having interpreting formulas that describe coordinate transformations. The target level of teacher effort chosen is approximately one-minute per student post. The particular facet cluster of interest is "Coordinate transformation with reflection." Pregnant posts for such a cluster can be elicited with specific questions on activity sheets, and this was done in this example. The following question was posed.

"Suppose we were to load the Mona Lisa image (actually mona-rgb.jpg) and use it as Source1. And suppose we created a new image that was twice as wide (846) and just as high (421) as the Mona Lisa image, and used it as Destination. What would happen if we applied the formula below?

if $x < 423$ then Source1(x, y) else Source1$(845-x, y)$

Reply to this message with your prediction."

One student posted the message

"I think that if you apply that formula to the mona lisa, the image will rotate itself, probably to 180 degrees, we'll find out."

The selected portion of this message that serves as evidence of a facet is "will rotate itself, probably to 180 degrees". This concisely expresses the student's notion that the formula represents a rotation. The facet that best corresponds to this is "confuses a rotation with a reflection." Creating a facet-assessment record is a formality within INFACT and is done by simply clicking on a button. Associating a certainty value with the record permits the teacher to record a judgment about the likely accuracy of the diagnosis. A reasonable value here is 4, because the student's prediction fits the (inexpert) facet well, but the teacher could interpret the last part of the message, "we'll find out" to be an expression of the student's own doubt and therefore some degree of disbelief in this facet.

4.4. Interventions

Regardless of the complex landscape of a student's understanding, the ultimate goal is to decide upon an appropriate intervention to help the student progress towards some learning objective. When a teacher has made a diagnosis, how does s/he determine what interventions are appropriate, and for whom? In general, this is still an open question.

Some of the possible actions a teacher can take are (a) to suggest that a student having a problematical facet pair up with a student having the expert-like facet for an explanation, (b) to suggest a particular piece of reading, particular problem to work, or web page to go to, or (c) arrange for individual instruction from the teacher or a teaching assistant. In the future, computer-based intelligent tutors may be able to offer facet-specific interventions that are designed to help students improve their understanding of the concept as efficiently as possible.

For the image-processing example, a suitable intervention for this student would consist of (a) encouraging the student to go ahead and use the computer to apply the formula to the image and see the result, and (b) calling attention to the part of the formula that represents the reflection and suggesting that the student actually plug in test values for x and y and see what the formula does with them.

5   Discussion

In this paper, we have explained the reasons for using facets, and we've described a process for developing a catalogue of facets. If this methodology were used more widely, an important question would be, Should we try to have standard facet catalogues for each given curriculum or even each given subject? One advantage to having one standard facetbase for, say, calculus, would be that any diagnoses made with one tool (e.g., Diagnoser) would be transferable to another system (e.g., INFACT), and so the various features of different tools can complement each other in a given educational environment.

On the other hand, standardization has its problems. Facets can be expected to change as the student population changes over the years; student preconceptions are partly shaped by their cultural experience, by the media, and by the activities in which they have taken part in the past. Facetbases will need to be updated over time, meaning that there will inevitably be different versions of each facetbase. Not only that, but many facets may be pedagogy-specific, and there will always be alternative opinions about the best way to teach a given subject. Thus reaching consensus on the facets in a facet catalogue may be difficult or too time-consuming in some situations.

Facet catalogues are related to ontologies as used in artificial intelligence, databases, and information retrieval. The Semantic Web is an example of a system where much consensus-building activity has led to standards for tagging material on the web. Ontologies have been a key idea behind the Semantic Web. There are good reasons to think about ontologies when designing facet catalogues, too. However, there are also practical reasons for avoiding some of the philosophical challenges of ontologies. Even if one embraces the notion of ontology for facet catalogue construction, one can still proceed without having to standardize; standardization presents so many challenges that alternatives are attractive. One alternative is to support transfer through ontological mappings as suggested by Tanimoto (2001).

Facet catalogues as we have described them are somewhat teacher-centric. They are created by master teachers or educational researchers, and they are created for use by teachers. The explanations of facets within them are to be read by teachers, not students. An important issue is the extent to which they can be made to directly serve students as well as teachers. One part of this issue is the possibility of keying feedback for students to the facets. In particular, can the explanations for teachers be slightly modified so as to provide explanations for students? The answer is probably yes; however, the explanations for facets in a facet catalogue are presumably de-contextualized – they represent conceptualizations somewhat apart from the particular activities and particular examples students may be working on. Therefore, the explanations probably have to be regenerated in terms of the student activities in which the facets are exhibited. In other words, the explanations must be made for the particular context in which students are working.

Future research and development work on facets includes (1) the creation of more flexible tools for diagnosis, with greater and greater degrees of automation, (2) the ability to diagnose facets from a wider and wider variety of data: online writing, digitized speech, sketches, log files from tools, and (3) better machine learning methods for facet diagnosis as suggested by Carlson and Tanimoto (2003).

6   References

Black, P. and D. Wiliam (1998). "Inside the Black Box: Raising Standards Through Formative Assessment." *Phi Delta Kappan* 80(2): 139-148.

Bransford, J. D., A. L. Brown, et al., Eds. (1999). *How People Learn: Brain, Mind, Experience and School*, National Academy Press.

Carlson, A., and S. Tanimoto (2003). Learning to identify student preconceptions from text, *Proc. HLT/NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, Edmonton.

diSessa, A. A. (1985). *Knowledge in Pieces*. Berkeley, University of California.

Hammer, D. (1996). "Misconceptions or p-prims. How might alternative perspectives of cognitive structure influence instructional perceptions and intentions." *Journal of the Learning Sciences* 5(2): 97-127.

Hunt, E. (c. 2000). Facet-based instruction. http://depts.washington.edu/huntlab/diagnoser/facet.html.

Hunt, E. and J. Minstrell (1994). A collaborative classroom for teaching conceptual physics. *Classroom lessons: Integrating cognitive theory and classroom practice.* K. McGilly. Cambridge, MIT Press.

Hunt, E. and J. Minstrell (1996). "Effective instruction in science and mathematics: Psychological principles and social constraints." *Issues in Education* 2(2): 123-162.

Hunt, E. and Pellegrino, J. W. (2002). Issues, examples, and challenges in formative assessment. *New directions for Teaching and Learning*, No. 89: 73-85.

McCloskey, M. (1983). Naive theories of motion. *Mental models*. D. Gentner, Stevens, A. L. Hillsdale and London, Lawrence Erlbaum: 299-324.

Minstrell, J. (1992). Facets of students' knowledge and relevant instruction. *Research in physics learning: Theoretical issues and empirical studies*. R. Duit, F. Goldberg and H. Niedderer. Kiel, IPN: 110-128.

Minstrell, J. (2000). Student Thinking and Related Assessment: Creating a Facet Assessment-based Learning Environment. *Grading the Nation's Report Card: Research from the Evaluation of NAEP*. J. Pellegrino, Jones, L., Mitchell, K. Washington DC, National Academy Press.

Minstrell, J. (2001). Facets of students' thinking: Designing to cross the gap from research to standards-based practice. *Designing for Science: Implications for Professional, Instructional, and Everyday Science*. K. Crowley, C. D. Schunn and T. Okada. Mahwah, Lawrence Erlbaum Associates.

Redish, E. F. (2004). A Theoretical Framework for Physics Education Research: Modeling student thinking. *Proceedings of the International School of Physics*, "Enrico Fermi" Course, IOS Press.

Scalise, K. and M. Wilson (2005). "Bundle Models for Data Driven Content in E-Learning and CBT: The BEAR CAT Approach."

Tanimoto, S. (2001). Distributed transcripts for online learning: design issues. *Journal of Interactive Multimedia in Education*. [www-jime.open.ac.uk/2001/2]. Publ. 10 Sept., 2001. ISSN:1365-893X.

Tanimoto, S., A. Carlson, J. Husted, E. Hunt, J. Larsson, D. Madigan, and J. Minstrell (2002). Text forum features for small group discussions with facet-based pedagogy, *Proc. CSCL 2002*, Boulder, CO.